SSOAR

Open Access Repository
www.ssoar.info

Measuring attitudes towards immigration across countries with the ESS: potential problems of equivalence

Rother, Nina

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Rother, N. (2005). Measuring attitudes towards immigration across countries with the ESS: potential problems of equivalence. In J. H. P. Hoffmeyer-Zlotnik, & J. Harkness (Eds.), *Methodological aspects in cross-national research* (pp. 109-125). Mannheim: GESIS-ZUMA. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49154-7

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

MEASURING ATTITUDES TOWARDS IMMIGRATION ACROSS COUNTRIES WITH THE ESS

Potential Problems of Equivalence

NINA ROTHER

This paper analyses the functional equivalence of attitudes towards immigration in internationally comparative research. In order to provide a knowledge base on attitudes towards immigration in Europe, a special module was included in the ESS round 1. However, attitudes towards immigration are not easy to measure and especially difficult to compare across cultures as the underlying concepts of immigration and immigrants may vary across countries.

Therefore a cross-cultural comparison of attitudes towards immigration can only be done when functional equivalence is given. Functional equivalence of data can be affected by three kinds of bias: construct bias, method bias and item bias. In order to analyse the functional equivalence of ESS immigration items, a secondary analysis of the ESS data was conducted.

The results show that some of the ESS measures might not be regarded as functionally equivalent. Some possible solutions are outlined.

1 Introduction

To know about attitudes towards immigration is essential, not only for researchers, but especially for politicians. Immigration plays a more and more important role in European societies nowadays, as can be seen in growing concerns about refugee issues or the importance to lower restrictions for highly skilled migrants that are needed in certain industries. If we just think of recent events as the EU enlargement or the Cap Anamur incident, followed by discussions about refugee camps in North Africa, it is getting clear how important it has become again to know more about what "attitudes towards immigration" are like in European states.

In order to provide researchers with knowledge and a database on attitudes towards immigration within Europe, the European Social Survey (ESS) round 1 contains a special module on attitudes towards immigration. So the ESS seems to be ideal for answering those questions on what those attitudes are like in Europe. However, it is not easy to measure, and eve more difficult to compare attitudes towards immigration across cultures. Different migration histories and policies in the different European countries make it hard to think of a common understanding of what immigration and immigrants are. Therefore, before starting with substantive analyses, it should always be proved that the data are comparable, i.e. that they are functionally equivalent.

After a short discussion of functional equivalence, types of biases and possible detection methods, a few potential biases found within the ESS data on attitudes towards immigration will be shown on an exemplary basis.

2 Bias and Equivalence

Lots of bias definitions and huge error lists from different research areas such as cross-cultural survey research (Braun, 2003) or cross-cultural psychology can be found (Johnson, 1998; Hui & Triandis, 1985; Berry et al., 1993; van de Vijver, 1998). For the purpose of this paper, it is especially relevant to focus on the psychological perspective of equivalence. For cross-cultural psychologists, data are equivalent if the underlying theoretical concepts, as well as the scores, have a similar meaning in the different countries. The negative counterpart of equivalence – bias – is used as a generic term to describe all nuisance factors that threaten the validity of cross-cultural comparisons (van de Vijver & Leung, 1997). Following the classification of van de Vijver & Tanzer (1997), cross-cultural psychologists further distinguish between three levels of bias: construct, method and item bias.

We are talking of *construct bias* if a not identical construct is measured across cultural groups. van de Vijver & Poortinga (1997) give an overview of possible causes of construct bias. They mention, for example, an incomplete overlap of definitions or an incomplete coverage of the construct – which is also called construct underrepresentation (Embretson, 1983). Also, a poor sampling of all relevant behaviours or a differential appropriateness of the questionnaire content can cause a construct bias. In the present case of attitudes towards immigration in the ESS, a construct bias due to an incomplete overlap of definitions could be the case if the questionnaire does not specify the term “people who come to live here” so that respondents in some countries would include refugees while respondents in other countries would not include them because of a different composition of immigrants in those countries.

In contrast to method and item bias, a construct bias affects the most basic level of data: the *structural* comparability of data, which comes first, i.e. before the comparison of the level of attitudes. So it always has to be analysed first if structural equivalence is given. Only then, level-oriented analyses are reasonable. Such level-oriented analyses can then still be affected by method and item bias.

The standard way to detect a "construct bias" is to compute factor analyses and then to look for differences in the factor solutions (van de Vijver, 2003). Differences in the factor solutions of different countries then indicate a construct bias. Two different kinds of factor analyses allow finding out about construct bias: On the one hand a confirmatory factor analysis for multi-groups can be computed (see for example Bollen, 1989), which allows various tests of model fit but does not offer an index of factorial agreement.

On the other hand an exploratory factor analysis using *target rotation*, also know as Procrustes Rotation, can also help to find out if structural equivalence is given (Harman, 1976). A factor analysis has to be computed in which the country factor loadings are rotated towards the general solution in order to correct for the arbitrariness of rotations in factor analyses. After that, as an indicator of factorial agreement, different coefficients of agreement like the identity coefficient, the additivity coefficient or the proportionality coefficient, which is also known as *Tucker's phi* can be calculated (van de Vijver & Leung, 1997). As the latter coefficient is insensitive to multiplications but influenced by additions, it seems to be the appropriate coefficient for comparisons of the structural equivalence of the immigration dataset.

But apart from factor analysis, it is always wise to collect additional data on the culture-level to investigate the underlying construct in detail, possibly by conducting cognitive interviews. It is only by getting more information on the different function of the biased items in the different countries, that you can find out what the real reasons for the construct bias have been and how the data can be corrected.

The second type of bias, *method bias*, is a bias that occurs within the scores due to particular characteristics of the instrument or the administration. It therefore affects scores at the level of the whole instrument (van de Vijver & Leung, 1997). Possible causes are many and various, such as a differential response style – like social desirability for example –, interviewer effects or communication problems between the respondent and the interviewer. The presence of a method bias can be seen through significant cross-country differences. However, a mixture of valid cross-country differences and method bias cannot be excluded so that it is wise to utilise other methods such as cognitive tests and monotrait-multimethod designs here as well.

Finally, we are talking of an *item bias* if persons from different cultures, but with the same standing on the underlying construct have a different score on the item (van de Vijver & Leung, 1997). Causes of item bias are measurement artefacts at the item level such as a poor item translation, an inadequate item formulation or differences in the appropriateness of the item content. In order to detect an item bias, a reliability analysis, for example by comparing Cronbach's α, and all exploratory methods, such as a comparison of means, Item-Response-Theory, Multidimensional Scaling or ANOVA can be used. But nevertheless, a mixture of valid cross-country differences and item bias again can occur, so that cognitive interviews to find out how the items are understood in different countries, might be useful here as well.

Although method bias definitely is important, I will now only discuss and analyse the presence of construct and item bias in the ESS data on immigration.

3 Dataset, Variables and Countries

The database for my analyses was the immigration module (D) of ESS round 1, edition 4.1. In principle, there are 58 questions on topics related to different immigration phenomena that could be analysed. However, all items on attitudes towards refugees (D49-D55) were excluded because I did not want to include attitudes towards a subgroup of immigrants but only general attitudes. All items that measure perceptions or estimations rather than attitudes were excluded, too. In detail, the items D1-D3, D38-D39, D47-D58 were excluded and 39 items on "attitudes towards immigration" remained for further analyses.

As for the countries to analyse, 20 countries will be compared and constitute the basis for all of the following analyses: Austria, Belgium, the Czech Republic, Denmark, Finland, Germany, Greece, Hungary, Italy, Ireland, Luxembourg, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland and the United Kingdom[1].

For the analyses of construct and item bias in this paper, however, I will focus on 5 countries which can serve as examples of the 20 countries. Germany and the UK represent the *classic immigration countries*. In these two countries, immigration has a very long history beginning with migrants from the former colonies immigrating to the UK and Polish mine workers settling in the German Ruhr basin at the beginning of the 20th century. In the early 1970ies both countries filled their labour shortages with workers mainly from Southern Europe. Nowadays a lot of refugees and asylum seekers still arrive in both countries as well as other migrants who come for reasons of family reunification. As *new immigra-*

1 For my analyses, there were 20 rather than 21 countries as I decided to exclude Israel from the beginning. Israel is simply too far away from the European context especially with regard to its immigration history.

tion countries, Italy and Portugal were chosen. Both countries have been emigration countries for a long time, sending lots of their inhabitants to Northern Europe. Only in the last few years, a back-migration has started. Additionally, both countries now have to deal with the growing problem of illegal immigrants that arrive on the shores of both countries. Finally Luxemburg, with its multicultural and multilingual society serves as a very interesting *special case*. The OECD data for 2000 (OECD, 2003) show that the percentage of foreigners in Luxembourg with about 37% lies well above the EU average. The rate of EU foreigners with almost 78% is even more impressive. Apart from this remarkable composition of the population, Luxembourg has three official languages: French, Luxembourgish and German and is therefore unique, regarding its societal structure.

4 Analysis

Before analysing the data regarding possible biases and therefore a functional inequivalence, some preliminary analyses had to be conducted.

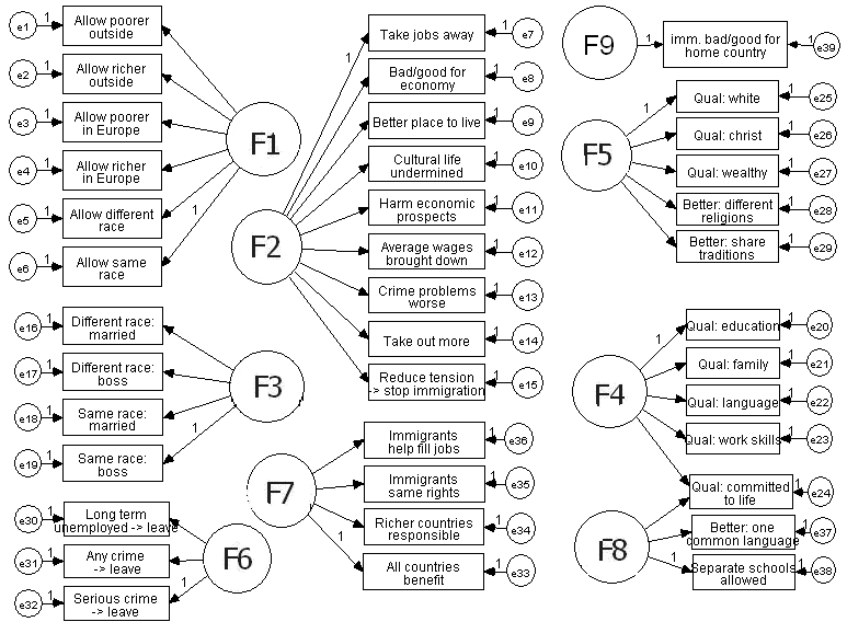
4.1 Preliminary analysis: latent structure of variables

First of all, the underlying factor structure of the 39 items had to be found out. This factor structure should assign all of the 39 items to a smaller number of factors. Furthermore, this factor structure should be a general one and therefore apply to all 20 countries in the same way.

As Christian Dustman, Ian Preston and their colleagues, who developed this immigration module for the ESS, have not yet postulated a concrete structure of their 58 items, I had to create such a structure myself by looking at the items and the theory. The result, after conducting various factor analyses and comparing the results, was a model of 9 factors. These 9 factors can be interpreted as follows:

- Factor 1 measures attitudes towards the *quantity of immigrants* that should be allowed to come into the country. It consists of six items, namely D4-D9.
- Factor 2 comprises attitudes towards the *effects of immigration on the host country*. For example, the item "immigrants take jobs away" (D25) is included here, as well as D18-D19, D26-30 and D44.
- Factor 3 captures which feelings people have about foreigners in their *personal environment*. Four items, D34-37, indicate how much one would mind or not mind if a person having either the same race or a different race would be appointed as one's boss or married a close relative.
- Factor 4 – *qualification* – indicates, how important it is that immigrants bring with them different qualifications such as a high education, a close family in the host country or the ability to speak the host country's language (D10-D12, D16).

Figure 1 Underlying Factor Structure of the 38 Immigration Items



Structural Equivalence of overall Factor Structure

- Factor 5 indicates which *cultural characteristics* like "being white" (D14), "coming from a Christian background" (D13) or "being wealthy" (D15) are important for immigrants to come. However this factor also comprises two items, which indicate how much a culturally homogenous society is desired by the respondents (D40-D41).
- Factor 6 measures *attitudes towards a deportation policy* and indicates under which circumstances people favour that immigrants are made to leave the country. This factor comprises the items D21 and D23-D24.
- Factor 7 holds a *general evaluation of immigration*, for example whether immigrants should get the same rights as everyone else (D22). Apart from this item, D20 as well as D32-33 also belong to this factor.
- Factor 8 captures attitudes towards a *cultural integration* of immigrants. Three items show high loadings on this factor, namely "better for a country if there is one common language" (D 42), "separate schools should be allowed" (D43) and "immigrants should be committed to the way of life" (D17).

- Factor 9 measures the perceived *effects of immigration on the immigrants' home country*. As it only consists of the item "immigration good/bad for those countries in the long run" (D31), this factor will be excluded from further analyses so that 8 factors remain.

Figure 1 shows the visualisation of the underlying factor structure, which could also be used for analysing it using confirmatory factor analysis.

4.2 Structural equivalence of overall factor structure

Before the data can be analysed on the substantive level by comparing the factor scores across countries, it has to be proved that the underlying factor structure applied to all of the analysed countries, to make at least sure that no construct bias has occurred. Therefore, an exploratory *principal component analysis with target rotation* is computed to find out if the same underlying factor structure can be assumed in the five selected countries. This is done by computing a principal component analysis in which the country factor loadings were rotated towards the general solution, that is the factor loadings of all 20 countries taken together. Tucker's phi was calculated afterwards as indicator of factorial agreement.

Table 1 shows the results for the 8 factors in the 5 countries. According to the suggestions of van de Vijver & Leung (1997), coefficients smaller or equal .85 have to be interpreted by all means as an indicator of missing factorial agreement. Thus factors 1 to 4 and also factor 7 show high values of factorial agreement in all countries. So it seems that the country factor structures correspond to the general assumed factor structure and no construct bias is given for these factors. In Great Britain, this also applies to all factors. Therefore it can be assumed that the British factor solution is the same as the general one and structural equivalence is given in Great Britain. This might be true because the questionnaire was developed in English, by British researchers who probably know best the situation in their home country and also have experience as to how questions are interpreted and answered in Great Britain. In Great Britain, the whole questionnaire was also pretested and possible problems thus could have been eliminated before the fieldwork started[2].

2 The second country where the items have been pretested were the Netherlands. It is a very interesting result that also in this country no deviations from the general factor structures could be found. Therefore, it seems that pretesting and adapting the item formulation according to the results of the pretests, is quite an important part of the questionnaire development, leads to big improvements of the data quality and can avoid time-consuming analyses.

Table 1 Values of Factorial Agreement Using Tucker's phi

	F1	F2	F3	F4	F5	F6	F7	F8
Germany	.99	.99	.99	.99	.98	.85	.94	.88
UK	.99	.99	.99	.99	.98	.96	.96	.96
Italy	.98	.98	.98	.94	.63	.94	.89	.82
Luxembourg	.98	.96	.96	.92	.82	.78	.91	.88
Portugal	.97	.96	.97	.95	.87	.92	.90	.53

Source: ESS round 1, edition 4.1; weighted with dweight; values equal or below .85 in italics

For the four countries other than Great Britain, the values are sub-optimal. Regarding factor 5, Italy and Luxembourg are problematic as they have quite small coefficients with a Tucker's phi of only.63 in Italy and .82 in Luxembourg. The same is true for Germany and Luxemburg for factor 6, for which Germany shows a Tucker's phi of .85 and Luxembourg of .78. Italy also has a small value of factorial agreement for factor 8 (phi = .82) as well as Portugal (phi = .53). So it has to be assumed that it has to be dealt with a construct bias in these four countries for these three factors.

A low factorial agreement in principle can have two reasons. Firstly, the loadings of the items that belong to the biased factor have evidently higher or lower loadings on this factor than in the general factor solution. The second possible reason for a low factorial agreement of a factor is that other items, which in fact do not belong to the biased factor, show too high loadings on it.

In the following analyses, it will be checked for factor 6 and factor 5 which of the two possible reasons of construct bias can be found. Furthermore, it will be checked what the reasons for these deviations could be and if they are due to an item bias of some of the items.

4.3 Structural equivalence of factor 6 – "deportation policy"

Factor 6, measuring attitudes towards a deportation policy, showed a missing factorial agreement in Germany and Luxembourg. Possible reasons for this construct bias will now be looked at, firstly by analysing the factor loading matrices, then by going down on the item level and comparing item means and finally by checking the linguistic equivalence of the factor 6 items.

Comparison of factor loadings

The factor loadings of factor 6 and other items on the biased factor 6 are given in Table 2. In the upper part of this table, you can check if the loadings of the three factor 6 items of the two biased countries show too high or too small values by comparing them with the loadings of these items of the general solution.

Table 2 Factor Loadings of Selected Items on Factor 6

		All countries	GER	LUX
F6	Serious crime → made to leave	.81	*.55*	*.51*
	Any crime → made to leave	.78	*.58*	*.66*
	Long term unemployed → made to leave	.50	.47	.43
F2	Immigration bad or good for country's economy	.07	.26	.19
	Taxes + services: immigrants take out more/less than put in	.07	*.28*	*.24*
	Immigrants make country worse / better place to live	.12	*.31*	.21
	Average wages/salaries generally brought down	.12	-.08	*-.43*
	Country's crime problems worse / better	.18	.31	*.43*
F7	Richer countries responsible to accept people	-.08	-.07	*.16*
	All countries benefit if people can move where skills needed	.06	*.40*	.18
	Immigrants help to fill jobs where there are shortages	.02	*.25*	-.01

Source: ESS round 1, edition 4.1; weighted with dweight; factor loadings higher than .2 and a deviation of more than .15 from the general loadings in italics.

It can be seen, that in both biased countries, the 3 items belonging to factor 6 have lower loadings than in the general solution, especially the items "serious crime → made to leave" and "any crime → made to leave". Therefore, it can be assumed that in Germany and Luxembourg the impact of attitudes on deportation in case of criminal behaviour on this factor is much weaker in these two countries than in the general solution. So in these two countries, factor 6 measures less attitudes towards deportation in case of criminal behaviour but rather a general evaluation of deportation of foreigners.

Another possible reason for a low factorial agreement of a factor could be, as discussed above, too high factor loadings of items that do not belong to the biased factor 6. In the lower part of Table 2, you can see that in Germany and Luxembourg variables that should only belong to factor 2 and 7 have quite high loadings on factor 6. In Germany this is true for the items "immigration bad / good for country's economy", "taxes + services: immigrants take out more / less than put in", "immigrants make country worse / better place to live", "all countries benefit if people can move where skills needed" and "Immigrants help to fill jobs where there are shortages". In Luxembourg, the items "taxes + services: immigrants take out more / less than put in", "average wages/salaries generally brought down", "country's crime problems worse / better" and "richer countries responsible to

accept people" have too high loadings on factor 6[3]. As factor 2 measures general attitudes towards the effects of immigration on the host country and factor 7 captures a general evaluation of immigration, it seems that in Germany and Luxembourg factor 6 also has a tendency into this direction. It can therefore be assumed that in Germany and Luxembourg, factor 6 measures also only general attitudes towards immigration and not particular attitudes towards deportation policy as it does in the other countries.

The question remains, why the factor loading structure in Germany and Luxembourg is so different from the rest. In order to find out about this, a change of the abstraction level is indicated, Therefore, analyses on the item level are conducted to find out if a single item could be responsible for the construct bias by functioning differently in these countries.

Comparison of means

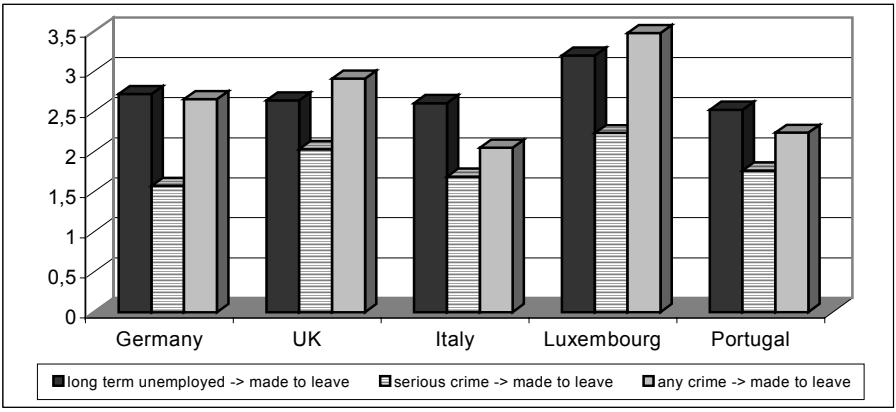
Firstly, a comparison of the means of the three factor 6 items across the five countries can show differences that are not only due to cross-country attitude differences but that are due to a differential functioning of a single item and therefore have to be interpreted as an item bias which then again could have affected the construct equivalence (van de Vijver & Leung, 1997).

In Figure 2 the means of the factor 6 items can be compared. Low means stand for an agreement on the respective item, i.e. immigrants should be made to leave, and high means indicate a tolerant attitude, i.e. immigrants should not be made to leave. The answering scale ranges from 1 to 5. It can be seen that in all countries, the agreement on a deportation of immigrants is highest in case of a serious crime. You can also see that out of all five countries, Luxembourg is the most tolerant one, having the highest means. But it also turns out that in Germany and Luxembourg, the differences between the item "serious crime" and the two others is bigger than in the other three countries[4]. People from Germany and Luxembourg are, in comparison to the UK, Italy and Portugal, much more in favour of making immigrants leave their countries after they have committed a serious crime, than if they only had committed "any crime" or had been "long term unemployed".

3 As too high loadings, I consider factor loadings that are a) higher than .2 and b) deviate with more than .15 from the general loadings.

4 This can be seen more clearly by comparing the relative difference of the mean scores across countries. The ratio of the average of the items "unemployed → made to leave" and "any crime → made to leave" to the item "serious crime → made to leave" is 1.70 in Germany and 1.50 in Luxembourg and therefore higher than in the three other countries where the ratio is always lower than 1.38.

Figure 2 Mean Scores of Factor 6 Items across Countries



Source: ESS round 1, edition 4.1; weighted with dweight

Two reasons for this deviation are possible. Either the true values really differ in Germany and Luxembourg or an item bias has occurred. If the latter is true, then the construct bias of the whole factor also becomes understandable: As the item "serious crime" works differently in Germany and Luxembourg, it also has a lower loading and therefore less impact on factor 6 which then measures more 'general attitudes' than 'detailed attitudes' towards deportation in case of criminal behaviour. Even though it would be easier to find out we have to deal with an item bias if cognitive interviews were available, it is still possible to find explanations for an item bias by taking a look at the society for example.

In Germany, in the course of the discussions about the new immigration law, deportation policy was discussed intensively in the media. Germans also remembered the trouble that the Bavarian government had when they had wanted to send a 14-year old Turkish serial delinquent back to Turkey. Similar reasons might also apply to multicultural Luxembourg that seems to be quite sensitive towards everything concerning safety policy.

Another reason for an item bias of the "serious crime" item and consequently also the construct bias of the whole factor 6 in Germany and Luxembourg might be found by checking the item formulation and translation itself.

Check of linguistic equivalence

The English source questionnaire uses the wording "If people who have come to live here commit a serious crime, they should be made to leave" which can easily be understood as meaning an active deportation by the government.

But in German, the item formulation is "Wenn Zuwanderer, die nach Deutschland kommen, eine schwere Straftat begehen, dann müssen sie das Land wieder verlassen." This is much weaker and only has the meaning of "delinquents have to or must leave the country". Here it remains open how they should leave – if the government has to deport them actively or if the delinquents themselves should be reasonable enough to leave the country. This difference in meaning also applies to the other 2 factor-6-items. So once you realise that the German questionnaire did not ask explicitly for the people's opinion on an active deportation by the government, it becomes understandable that factor 6 has much in common with factor 2 and 7, both of which measure attitudes towards immigration on a more general level.

In the translation for Luxembourg, "Si les gens, qui sont venus vivre ici, commettent une infraction grave, il faut les obliger à repartir.", there does not seem to be any semantic differences to the wording in the source questionnaire, so it seems that the item is linguistically equivalent.

To find out which item bias possibly occurred in Luxembourg and to confirm the hypothesis about the reasons for the construct bias in Germany, some cognitive interviews are needed.

4.4 Structural equivalence of factor 5 – "cultural characteristics"

As Table 1 showed, it is Italy and again Luxembourg that have a low score of factorial agreement on factor 5, which measures attitudes towards cultural characteristics of immigrants. Possible reasons for this construct bias will now be looked at, using the same methods as for the analysis of the construct bias of factor 6.

Comparison of factor loadings

The upper part of Table 3 again shows the factor loadings of the five items associated with factor 5 on that factor. It can be seen that – especially in Italy – there are quite big deviations from the general factor solution. All of the five items show smaller loadings on factor 5. This is also true for two items in Luxembourg, namely the item "Qualification for immigration: be white" and "Better for country if almost everyone share customs and traditions".

Table 3 Factor Loadings of Selected Items on Factor 5

		All countries	ITA	LUX
F5	Better for a country if a variety of different religions	.50	*.28*	.54
	Qualification for immigration: Christian background	.67	*.39*	.65
	Qualification for immigration: be white	.69	*.42*	*.45*
	Better for a country if almost everyone share customs and traditions	-.50	*-.31*	*-.05*
	Qualification for immigration: be wealthy	.56	*.28*	.43
F4	Qualification for immigration: good educational qualifications	.05	*.32*	.12
	Qualification for immigration: speak country's official language	.11	*.36*	.09
	Qualification for immigration: close family living here	.11	*.32*	.22
F3	Immigrant different race/ethnic group majority: your boss	.17	.17	*.37*
	Immigrant different race/ethnic group majority: married close relative	.19	.15	*.46*
F8	Better for a country if almost everyone speaks one common language	-.01	*-.21*	*.21*

Source: ESS round 1, edition 4.1; weighted with dweight; factor loadings higher than .2 and a deviation of more than .15 from the general loadings in italics.

The lower part of Table 3 shows which of the other items show – misleadingly – high loadings on factor 5 in Italy and Luxembourg. In Italy, almost all of the items that in fact should belong to factor 4 show the same high loadings on factor 5 as the factor 5-items. In detail, this applies to the items "Qualification for immigration: good educational qualifications", "Qualification for immigration: speak country's official language", "Qualification for immigration: close family living here" and also the factor 8 item "Better for a country if almost everyone speaks one common language". A hypothesis could then be set up that for Italy no distinction between the qualifications and cultural characteristics can be made. Consequently, this would indicate a construct bias.

This different structure in Italy can also be seen in a confirmatory factor analysis in which covariances of the factors are allowed. Factor 4 and 5 correlate with almost one whereas the covariance in the general solution is rather high with .63, but significantly lower than 1.0. So it seems that in Italy factor 4 and 5 only constitute one single factor in reality. This combination then measures what Italians think immigrants should bring with them and how they should behave.

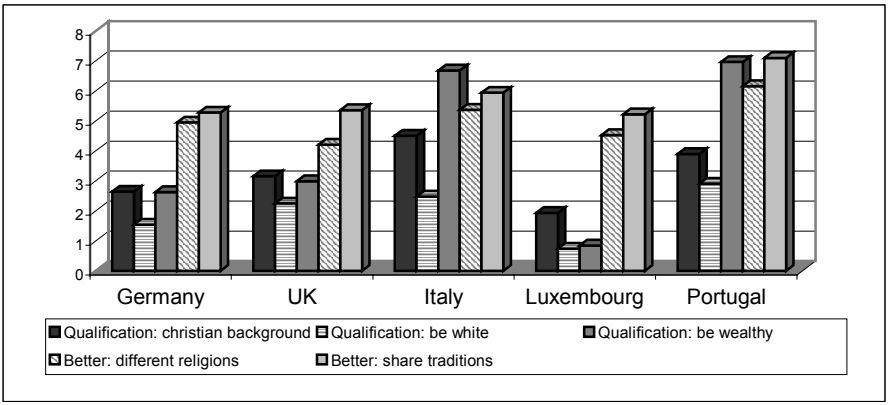
In Luxembourg, other items also have higher loadings on factor 5. These are the three items indicated in the last three lines of Table 3: "Immigrant different race/ethnic group majority: your boss", "Immigrant different race/ethnic group majority: married close relative" and "Better for a country if almost everyone speaks one common language". All of these items measure attitudes in relation to "culture", i.e. how important cultural char-

acteristics and ethnic background for personal contacts are. So, whereas in Italy, factor 4 and 5 form one general factor, it seems that in Luxembourg, factor 5 together with factor 8 and factor 3 constitute one big "culture factor". The hypothesis then is that, as Luxembourg is per se a multilingual country and also a country with a high proportion of highly qualified foreigners, problems of any kind of cultural integration and background are less relevant in social life. They are therefore less separated in people's minds so that no distinction into three distinct factors can be made in Luxembourg.

Comparison of means

A comparison of the mean scores of the factor 5 items in Figure 3 can again show interesting deviations across countries that could have an effect on the construct bias of factor 5 in Italy and Luxembourg. All items were recoded if necessary so that low values indicate a rather tolerant attitude, i.e. that it is not important that immigrants are wealthy, white etc. High means, on the other hand, stand for a less tolerant attitude, i.e. that it is very important that immigrants are wealthy, white etc. In order to facilitate the comparability, all of the items assessed on a 5-point-scale were transformed into a 11-point-scale ranging from 0 to 10.

Figure 3 Mean Scores of Factor 5 Items across Countries



Source: ESS round 1, edition 4.1; weighted with dweight

At first sight, it seems that Luxembourg is quite tolerant whereas Italy and Portugal are not. If the item "share customs and traditions" is compared, no big deviations across countries can be found, hence, no item bias seems to have occurred here. But a compari-

son of the item "be wealthy" shows that Italy and Portugal have very high means, whereas the mean in Luxembourg is very small. Here an item bias is possible. In Luxembourg, the item "be white" again has a very low mean, which could also point to an item bias. So an item bias should be checked for at each of these items. The case of Portugal should also be looked at more closely. Although the factorial agreement of this factor seems to be sufficient, there might be an item bias nonetheless.

Check of linguistic equivalence

As before, a check of the exact item formulations and translations is advisable in order to find out if the hypothesis of an item bias can be supported.

The English source questionnaire uses the following wording: "Please tell me how important you think each of these things should be in deciding whether someone born, brought up and living outside [country] should be able to come and live here. Firstly, how important should it be for them to…be white?" and "…be wealthy?". In Italy, the items were translated as "…avere la pelle bianca" and "avere una buona salute". The French translation for Luxembourg is "…avoir la peau blanche?" and "…être riche?". And finally, in Portugal, the items were translated into "…ser branco?" and "…ser rico?".

In Italian, the translation of "be white" is correct. But obviously there was a mix up: instead of "wealthy", the word "healthy" was translated in the Italian version. Taking this into account, the high mean becomes understandable and it is true that this item is biased and cannot be compared with other countries.

The French and Portuguese translation of "wealthy" into "rich" could also be problematic to some extent. It could be assumed that "wealthy" is likely to be interpreted as having lots of money and assets – which might not be considered as being that important for immigrants who have just arrived and want to make money. "Rich" on the other hand, might be rather considered as just having money, or having just enough money to earn one's living – something which is considered as important for immigrants.

But even if we assume that the French and Portuguese translation of "wealthy" into "rich" is correct, the items seem biased nonetheless. It is no wonder that people in Luxembourg, which is already a rich country and which predominantly receives fairly rich and white immigrants – mainly from within the EU –, say that money alone is not important for people to come and live in Luxembourg. The opposite is true for Portugal. Immigrants coming to Portugal mainly emigrate from former colonies like Cap Verde or Brazil as well as from Africa. They are predominantly poor and coloured immigrants. Taking these different contexts into account, it is risky to draw the conclusion that for people in Lux-

embourg colour does not matter and therefore that they are more tolerant than Portuguese who seem to be quite intolerant. If Luxembourg received poor, coloured immigrants, maybe the attitudes within the population would be different. Now, notwithstanding the fact that the wording is slightly different, there could still be an item bias here because with different social contexts, the item invokes different reactions in the two countries.

5 Conclusion

Two things should have been shown in this paper. First of all, it has to be stated that although a lot has been done in the ESS to avoid all types of bias, it seems impossible to avoid each and every single bias. On an exemplary basis, some of the few construct and item biases in the immigration module of the ESS were shown. It could also be seen that biases on the lower level, i.e. on the item level affect higher levels of equivalence, too. Therefore the great importance of the phase of questionnaire development was hopefully demonstrated.

Secondly, it was also shown, how important it is that every researcher interested in cross-country comparisons, first analyses if the data are equivalent for his purposes before starting substantive analyses.

Now, if any bias is found, there is no need to give up. If severe item biases occurred, the respective item should be excluded for the analysis in the respective country. This has already happened in edition 5 of the ESS with the Italian wealthy-healthy-item. For some countries a different factor structure will have to be assumed due to a construct bias. In the examples of this paper, this would have to be done for Luxembourg for instance, with its big culture factor or in the Italian case, with its combination of factor 4 and 5. If a different factor structure has to be assumed, the factor scores that are often the basis for further analyses should be calculated differently for each country by using confirmatory factor analysis and different models for each country.

In any case, additional data ought to be collected, for example by conducting cognitive interviews or by using split-ballot designs to make sure if or when a bias has to be dealt with and to find out why it has occurred.

References

Berry, J. W., Y. H. Poortinga, M. H. Segall, and P. R. Dasen. 1993. *Cross-cultural psychology. Research and Application.* Cambridge: Cambridge University Press.

Bollen, K. A. 1989. *Structural equations with latent variables.* New York: Wiley.

Braun, M. 2003. "Errors in Comparative Survey Research: An Overview." Pp. 137-142 in *Cross-Cultural Survey Methods,* edited by Harkness, J. A., F. J. R. van de Vijver, and P. Ph. Mohler. Hoboken: Wiley.

Embretson, S. E. 1983. "Construct validity: Construct representation versus nomothetic span." *Psychological Bulletin* 93:179-197.

Johnson, T. P. 1998. "Approaches to equivalence in cross-cultural and cross-national survey-research." Pp. 1-40 in *Cross-Cultural Survey Equivalence*, edited by J. A. Harkness. ZUMA-Nachrichten Spezial 3. Mannheim: ZUMA.

Harman, H. H. 1976. *Modern factor analysis.* Chicago: University of Chicago Press.

Hui, C. H., and H. C. Triandis. 1985. "Measurement in cross-cultural psychology. A review and comparison of strategies." *Journal of Cross-cultural Psychology* 16:131-152.

OECD. 2003. *Trends in International Migration. Annual Report, 2002 Edition.* Rome: OECD.

van de Vijver, F. J. R. 1998. "Towards a theory of equivalence and bias." Pp. 41-65 in *Cross-Cultural Survey Equivalence*, edited by J. A. Harkness. ZUMA-Nachrichten Spezial 3. Mannheim: ZUMA.

van de Vijver, F. J. R. 2003. "Bias and Substantive Analyses." Pp. 207-246 in *Cross-Cultural Survey Methods,* edited by Harkness, J. A., F. J. R. van de Vijver, and P. Ph. Mohler. Hoboken: Wiley.

van de Vijver, F. J. R., and K. Leung. 1997. *Methods and Data Analysis for Cross-Cultural Research.* Thousand Oaks: Sage.

van de Vijver, F. J. R., and Y. H. Poortinga. 1997. "Towards an integrated analysis of bias in cross-cultural assessment." *European Journal of Psychological Assessment* 13:21-29.

van de Vijver, F. J. R., and N. K. Tanzer. 1997. "Bias and Equivalence in Cross-Cultural Assessment: An Overview." *European Review of Applied Psychology* 41:263-279.